Understanding Honeypot Data by an Unsupervised Neural Visualization

Álvaro Alonso[1], Santiago Porras[1], Enaitz Ezpeleta[2], Ekhiotz Vergara[2], Ignacio Arenaza[2], Roberto Uribeetxeberria[2], Urko Zurutuza[2], Álvaro Herrero[1] and Emilio Corchado[3]

Abstract. Neural projection techniques can adaptively map high-dimensional data into a low-dimensional space, for the user-friendly visualization of data collected by different security tools. Such techniques are applied in this study for the visual inspection of honeypot data, which may be seen as a complementary network security tool that sheds light on internal data structures through visual inspection. Empirical verification of the proposed projection methods was performed in an experimental domain where data were captured from a honeypot network. Experiments showed that visual inspection of these data, contributes to easily gain a deep understanding of attack patterns and strategies.

Keywords. Projection Models, Artificial Neural Networks, Unsupervised Learning, Network & Computer Security, Intrusion Detection, Honeypots.

1 Introduction

A network attack or intrusion will inevitably violate one of the three computer security principles -availability, integrity and confidentiality- by exploiting certain vulnerabilities such as Denial of Service, Modification and Destruction [1]. One of the most harmful issues of attacks and intrusions, which increases the difficulty of protecting computer systems, is precisely the ever-changing nature of attack technologies and strategies.

[1] Civil Engineering Department, University of Burgos. C/ Francisco de Vitoria s/n, 09006 Burgos, Spain. {aad0038,ahcosio}@ubu.es

[2] Electronics and Computing Department, Mondragon University. Goiru Kalea, 2, 20500 Arrasate-Mondragon, Spain. {eezpeleta,evergara,iarenaza,ruruibeetxeberria,uzurutuza}@eps.mondragon.edu

[3] Departamento de Informática y Automática, University of Salamanca, Plaza de la Merced s/n, 37008 Salamanca, Spain. escorchado@usal.es

Visual inspection of traffic patterns is an alternative and crucial aspect in network monitoring [2]. Visualization is a critical issue in the computer network defence environment, which chiefly serves to generate a synthetic and intuitive representation of the current situation for the network manager; as a result, several research initiatives have recently applied information visualization to this challenging task [3] [4] [5] [6]. Visualization techniques typically aim to make the available statistics supplied by traffic-monitoring systems more understandable in an interactive way. They therefore focus on traffic data as well as on network topology. Regardless of their specific characteristics, these methods all map high-dimensional feature data into a low-dimensional space for presentation purposes. The baseline of the research presented in this study is that Artificial Neural Networks (ANNs), in general, and unsupervised connectionist models [7, 8], in particular, can prove quite adequate for the purpose of network data visualization through dimensionality reduction. As a result, unsupervised projection models are applied in the present research for the visualization and subsequent analysis of Honeypot data.

A honeypot has no authorised function or productive value within the corporate network other than to be explored, attacked or compromised [9]. Thus, a honeypot should not receive any traffic at all. Any connection attempt with a honeypot is then an attack or attempt to compromise the device or services that it is offering-is by default illegitimate traffic. From the security point of view, there is a great deal that may be learnt from a honeypot about a hacker's tools and methods in order to improve the protection of information systems.

In a honeynet, all the traffic received by the sensors is suspicious by default. Thus every packet should be considered as an attack or at least as a piece of a multi-step attack. Numerous studies propose the use of honeypots to detect automatic large scale attacks; honeyd [10] and nepenthes [11] among others. The first Internet traffic monitors known as Network Telescopes, Black Holes or Internet Sinks were presented by Moore *et al*. [12].

The remaining sections of this paper are structured as follows: section 2 presents the proposed approach and the neural projection techniques applied in this work. Some experimental results are presented and described in section 3; the conclusions of this study are discussed in section 4, as well as future work.

2 A Visualization-based Approach

This work proposes the application of projection models for the visualization of honeypot data. Visualisation techniques have been applied to massive datasets, such as those generated by honeynets, for many years. These techniques are considered a viable approach to information seeking, as humans are able to recognize different features and to detect anomalies by inspecting graphs [13]. The underlying operational assumption of the proposed approach is mainly grounded in the ability to render the high-dimensional traffic data in a consistent yet low-dimensional representation. So, security visualisation tools have to map high-

dimensional feature data into a low-dimensional space for presentation. One of the main assumptions of the research presented in this paper is that neural projection models will prove themselves to be satisfactory for the purpose of security data visualisation through dimensionality reduction.

This problem of identifying patterns that exist across dimensional boundaries in high dimensional datasets is a challenging task. Such patterns may become visible if changes are made to the spatial coordinates. However, an *a priori* decision as to which parameters will reveal most patterns requires prior knowledge of unknown patterns.

Projection methods project high-dimensional data points onto a lower dimensional space in order to identify "interesting" directions in terms of any specific index or projection. Having identified the most interesting projections, the data are then projected onto a lower dimensional subspace plotted in two or three dimensions, which makes it possible to examine the structure with the naked eye. Projection methods can be smart compression tools that map raw, high-dimensional data onto two or three dimensional spaces for subsequent graphical display. By doing so, the structure that is identified through a multivariable dataset may be visually analysed with greater ease.

Visualisation tools can therefore support security tasks in the following way:

- Visualisation tools may be understood intuitively (even by inexperienced staff) and require less configuration time than more conventional tools.

Providing an intuitive visualization of data allows inexperienced security staff to learn more about standard network behaviour, which is a key issue in ID [14]. The monitoring task can be then assigned to less experienced security staff.

As stated in [3], "*visualizations that depict patterns in massive amounts of data, and methods for interacting with those visualizations can help analysts prepare for unforeseen events*". Hence, such tools can also be used in security training.

They can work in unison with some other security tools in a complementary way.

As with other machine learning paradigms, an interesting facet of ANN learning is not just that the input patterns may be precisely learned/classified/identified, but that this learning can be generalised. Whereas learning takes place within a set of training patterns, an important property of the learning process is that the network can generalise its results on a set of test patterns that were not previously learnt. The identification of unknown patterns fits the 0-day attack [15] detection.

Due to the aforementioned reasons, the present study approaches the analysis of honeynet data from a visualization standpoint. That is, some neural projection techniques are applied for the visualization of such data. The different projection models applied in this study are described in the following sections.

2.1 Principal Component Analysis

Principal Component Analysis (PCA) is a standard statistical technique for compressing data; it can be shown to give the best linear compression of the data in terms of least mean square error. There are several ANNs or connectionist models which have been shown to perform PCA e.g. [16, 17, 18].

This technique describes the variation in a set of multivariate data in terms of a set of uncorrelated variables, in decreasing order of importance, each of which is a linear combination of the original variables. It should be noted that even if we are able to characterize the data with a few variables, it does not follow that an interpretation will ensue.

2.2 Cooperative Maximum Likelihood Hebbian Learning

The Cooperative Maximum Likelihood Hebbian Learning (CMLHL) model [19] extends the Maximum Likelihood Hebbian Learning (MLHL) [20] model, which is based on Exploratory Projection Pursuit (EPP) [21]. Considering an N- dimensional input vector (x), and an M-dimensional output vector (y), with W_{ij} being the weight (linking input j to output i), then CMLHL can be expressed as:

1. Feed-forward step: $y_i = \sum_{j=1}^{\mathbf{N}} W_{ij} x_j, \forall i$ (1)
2. Lateral activation passing: $y_i(t+1) = \left[y_i(t) + \tau(b - Ay) \right]^+$ (2)
3. Feedback step: $e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i, \forall j$ (3)
4. Weight change: $\Delta W_{ij} = \eta . y_i . sign(e_j) | e_j |^{p-1}$ (4)

Where: η is the learning rate, τ is the "strength" of the lateral connections, b the bias parameter, p a parameter related to the energy function [19, 20, 22] and A a symmetric matrix used to modify the response to the data [19]. The effect of this matrix is based on the relation between the distances separating the output neurons.

3 Experimental Study

Researchers usually make use of known attack datasets such as the well known DARPA dataset or the KDD Cup '99 subdataset in order to validate their devel-

oped systems. However, these data are simulated, non-validated and irregular [19] so they are not fully reliable. Even if the results obtained by such systems are good, no one can assure that the applied algorithms will make the system more secure or will detect real attacks. This is the main reason of using real attack data coming from a running honeynet in this work. The experimental work has been done by using data related to one month of real attacks that reached the 8 sensors used by the Euskalert project [23]. These data are depicted through different neural projections in order to discover real attack behaviour and strategies.

3.1 Euskalert Project

The Euskalert project has deployed a network of honeypots in the Basque Country (northern Spain) where eight companies and institutions have installed one of the project's sensors behind the firewalls of their corporate networks. The honeypot sensor transmits all the traffic received to a database via a secure communication channel. These partners can consult information relative to their sensor (after a login process) as well as general statistics in the project's website. Once the system is fully established, the information available can be used to analyse attacks suffered by the honeynet at network and application level. Euskalert is a distributed honeypot network based on a Honeynet GenIII architecture [21].

This honeypot system receives 4000 packets a month on average. All the traffic is analyzed by the Snort IDS, and an alert is launched whenever the packet matches a known attack signature. For this experiment, we have analysed the logs coming from Euskalert and Snort gathered during February 2010.

The February 2010 dataset contains a total of 3798 packets, including TCP, UDP and ICMP traffic received by the distributed honeypot sensors.

From this dataset, it may be said that a misuse detection-based IDS such as Snort is only capable of identifying about 10.38% of bad-intentioned traffic. Furthermore, it was demonstrated that only 2% of the unsolicited traffic was identified by the IDS when automatically generated signatures were included from a previous work [24]. Thus, a deeper analysis of the data is needed in order to discover the internal structure of the remaining 90% of the traffic. Explaining the behaviour of the unknown traffic is a difficult task that must be performed to better protect computer networks and systems.

3.2 Experiments and Results

The following features were extracted from each one of the records in the dataset:

- **Time**: the time when the attack was detected. Difference in relation to the first attack in the dataset (in minutes).
- **Protocol**: whether TCP, UDP or ICMP (codified as three binary features).

- **Ip_len**: number of bytes in the packet.
- **Source Port**: number of the port from which the source host sent the packet. In ICMP protocol, this represents the ICMP type field.
- **Destination Port**: destination host port number to which the packet is sent. In the ICMP protocol, this represents the ICMP type field.
- **Flags**: Control bits of a TCP packet, which contains 8 1 bit values.

The previously introduced projection techniques were applied to this dataset, generating the projections shown in this section. In these projections, the data are depicted with different colors and shapes, taking into account the different original features of the data. CMLHL was applied in order to analyse the dataset described above and to identify its inner structure.

Fig. 1 shows the CMLHL projection by considering the source port to depict the packets; from 3 to 10903: red circles, from 10904 to 21803: black crosses, from 21804 to 32703: green pluses, from 32704 to 43603: magenta stars, from 43604 to 54503: yellow squares, and from 54504 to 65401: cyan diamonds.

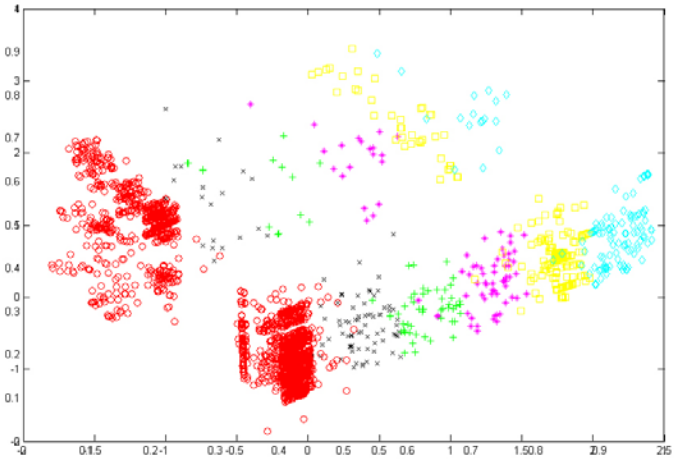


Fig. 1**.** CMLHL projection – source port**.**

Fig. 2 shows the CMLHL projection by considering the destination port to depict the packets; from 3 to 10371: red circles, from 10371 to 20739: black crosses, from 20739 to 31107: green pluses, from 31107 to 41475: magenta stars, from 41475 to 51843: yellow squares, and from 51843 to 62205: cyan diamonds.

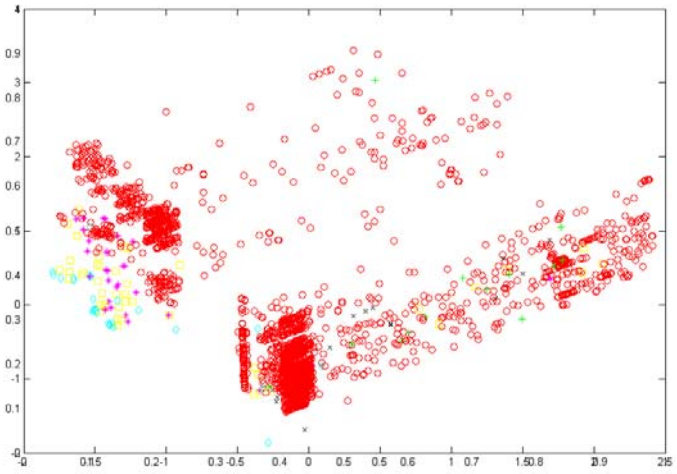


Fig. 2**.** CMLHL projection – destination port**.**

Fig. 3 shows the CMLHL projection by considering the time to depict the packets; from 0 to 6692: red circles, from 6693 to 13384: black crosses, from 13385 to 20076: green pluses, from 20077 to 26768: magenta stars, from 26769 to 33460: yellow squares, and from 33461 to 40148: cyan diamonds.

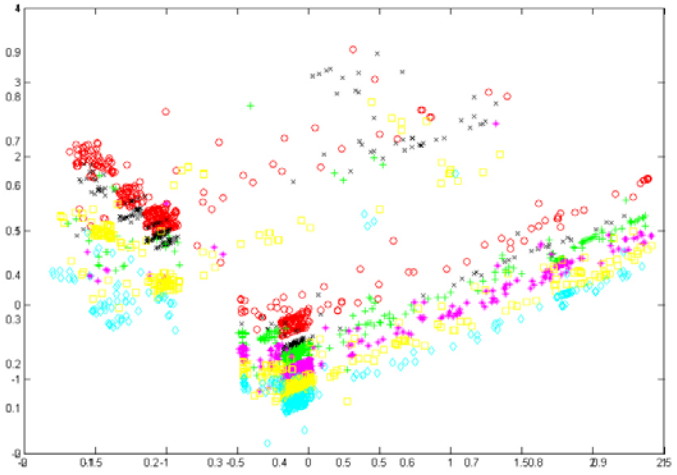


Fig. 3**.** CMLHL projection – time**.**

Fig. 4 shows the CMLHL projection by considering the protocol to depict the packets; ICMP: red circles, UDP: black crosses, TCP: green pluses.

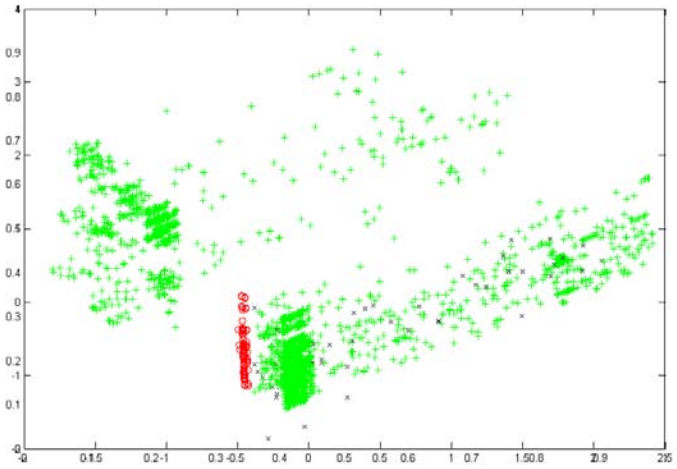
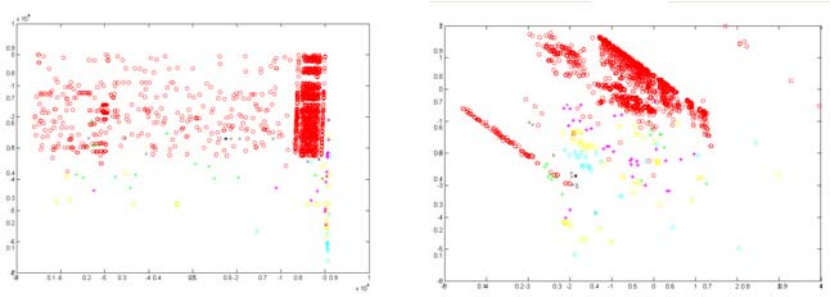


Fig. 4. CMLHL projection – protocol**.**

3.3 Comparative Study

For comparative purposes, some other projection techniques, namely PCA (see section 4.3.1) and MLHL, were applied to the previously described dataset. The obtained projections are shown in Fig. 5.



a) PCA projection – source port. b) MLHL projection – source port.

Fig. 5**.** Projections of data traffic captured by Euskalert, in February, 2010**.**

4 Conclusions and Future Work

From the projections in Figs. 1-5 we can conclude that CMLHL provides a more sparse representation that the other two methods. This enables the intuitive visualization of the honeynet, where the general structure of these data can be seen.

Thanks to the CMLHL projections it is easy to get a general idea of the dataset structure, an in-deep analysis can be subsequently carried out. For the analysed dataset, it can be concluded that visualizing attacks from a honeynet can help on identifying patterns that will inevitably end on home computers and companies. Such tactic can act as an early warning system to alert users before they can become victims of the observed attacks. Taking as an example the Figure 2, where destination port is depicted with CMLHL projections, every point but the red ones possibly indicate the backscatter phenomenon. The honeypot sensors received traffic coming from the real victim, as a response to a DoS or DDoS attack with spoofed source addresses, in this case the sensor itself. Looking at the red points, the highest density belongs to ports 135, 137 and 139. This are the most attacked services nowadays, corresponding to Windows vulnerable services.

Future work will combine the honeypot data with the output of a signature-based IDS, such as Snort, in the same visualization. This will validate the proposed approach as a complementary tool that can be combined with some other security tools or IDSs.

Acknowledgments. This research has been partially supported through the Regional Government of Castilla y León under Project BU006A08, the Regional Government of Gipuzkoa, the Department of Research, Education and Universities of the Basque Government; and the Spanish Ministry of Science and Innovation (MICINN) under project CIT-020000-2009-12 (funded by the European Regional Development Fund); project of the Spanish Ministry of Science and Innovation TIN2010-21272-C02-01 (funded by the European Regional Development Fund). The authors would also like to thank the vehicle interior manufacturer, Grupo Antolin Ingenieria S.A., within the framework of the MAGNO2008 - 1028.- CENIT Project also funded by the MICINN.

References

1. Myerson, J.M. (2002) Identifying Enterprise Network Vulnerabilities. International Journal of Network Management 12(3): 135-144
2. Becker, R.A., Eick, S.G., Wilks, A.R. (1995) Visualizing Network Data. IEEE Transactions on Visualization and Computer Graphics 1(1): 16-28
3. D'Amico, A.D., Goodall, J.R., Tesone, D.R., Kopylec, J.K. (2007) Visual Discovery in Computer Network Defense. IEEE Computer Graphics and Applications 27(5): 20-27
4. Goodall, J.R., Lutters, W.G., Rheingans, P., Komlodi, A. (2006) Focusing on Context in Network Traffic Analysis. IEEE Computer Graphics and Applications 26(2): 72-80
5. Itoh, T., Takakura, H., Sawada, A., Koyamada, K. (2006) Hierarchical Visualization of Network Intrusion Detection Data. IEEE Computer Graphics and Applications 26(2): 40-47

6. Livnat, Y., Agutter, J., Moon, S., Erbacher, R.F., Foresti, S. (2005) A Visualization Paradigm for Network Intrusion Detection. (ed) Sixth Annual IEEE SMC Information Assurance Workshop, 2005. IAW '05.
7. Herrero, Á., Corchado, E., Gastaldo, P., Zunino, R. (2009) Neural Projection Techniques for the Visual Inspection of Network Traffic. Neurocomputing 72(16-18): 3649-3658
8. Herrero, Á., Corchado, E., Pellicer, M.A., Abraham, A. (2009) MOVIH-IDS: A Mobile-Visualization Hybrid Intrusion Detection System. Neurocomputing 72(13-15): 2775-2784
9. Charles, K.A. (2004) Decoy Systems: A New Player in Network Security and Computer Incident Response. International Journal of Digital Evidence 2(3)
10. Provos, N. (2004) A Virtual Honeypot Framework. (ed) 13th USENIX Security Symposium 132.
11. Baecher, P., Koetter, M., Holz, T., Dornseif, M., Freiling, F. (2006) The Nepenthes Platform: An Efficient Approach to Collect Malware. (ed) 9th International Symposium on Recent Advances in Intrusion Detection (RAID 2006) 4219. Springer Berlin / Heidelberg. LNCS
12. Moore, D., Shannon, C., Brown, D.J., Voelker, G.M., Savage, S. (2006) Inferring Internet Denial-of-service Activity. ACM Transactions on Computer Systems 24(2): 115-139
13. Ahlberg, C., Shneiderman, B. (1999) Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. (ed) Readings in Information Visualization: using Vision to Think. Morgan Kaufmann Publishers Inc. 244-250
14. Goodall, J.R., Lutters, W.G., Rheingans, P., Komlodi, A. (2005) Preserving the Big Picture: Visual Network Traffic Analysis with TNV. (ed) IEEE Workshop on Visualization for Computer Security (VizSEC 05). IEEE Computer Society.
15. Laskov, P., Dussel, P., Schafer, C., Rieck, K. (2005) Learning Intrusion Detection: Supervised or Unsupervised? In: Roli, F., Vitulano, S. (ed) 13th International Conference on Image Analysis and Processing (ICIAP 2005) 3617. Springer, Heidelberg. LNCS
16. Oja, E. (1982) A Simplified Neuron Model as a Principal Component Analyzer. Journal of Mathematical Biology 15(3): 267-273
17. Sanger, D. (1989) Contribution Analysis: a Technique for Assigning Responsibilities to Hidden Units in Connectionist Networks. Connection Science 1(2): 115-138
18. Fyfe, C. (1997) A Neural Network for PCA and Beyond. Neural Processing Letters 6(1-2): 33-41
19. Corchado, E., Fyfe, C. (2003) Connectionist Techniques for the Identification and Suppression of Interfering Underlying Factors. International Journal of Pattern Recognition and Artificial Intelligence 17(8): 1447-1466
20. Fyfe, C., Corchado, E. (2002) Maximum Likelihood Hebbian Rules. (ed) 10th European Symposium on Artificial Neural Networks (ESANN 2002).
21. Friedman, J.H., Tukey, J.W. (1974) A Projection Pursuit Algorithm for Exploratory Data-Analysis. IEEE Transactions on Computers 23(9): 881-890
22. Corchado, E., MacDonald, D., Fyfe, C. (2004) Maximum and Minimum Likelihood Hebbian Learning for Exploratory Projection Pursuit. Data Mining and Knowledge Discovery 8(3): 203-225
23. McHugh, J. (2000). Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory. ACM Trans. Information System Security 3 (4), 262-294.
24. Zurutuza, U., Uribeetxeberria, R., Zamboni, D. (2008) A Data Mining Approach for Analysis of Worm Activity through Automatic Signature Generation. (ed) 1st ACM Workshop on AISec. ACM.